

1 Diberikan data sampel dibawah silahkan mencari dan menghitung mean (rata-rata), median, varians, standar deviasi. Gambarkan juga Histogram nya

72.2	31.9	26.5	29.1	27.3	8.6	22.3	89.2
20.4	12.8	25.1	19.2	24.1	58.2	68.1	55.4
55.1	9.4	14.5	13.9	20.7	17.9	8.5	
38.1	54.2	21.5	26.2	59.1	43.3	26.5	

a. Percentage of the families that are in the upper income level

- Rata - rata sampel atau mean

Rumus untuk mencari mean dari sample adalah sebagai berikut

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

dimana n adalah banyaknya data

Dan $x_1 + x_2 + \dots + x_n$ adalah jumlah data

Diketahui :

$$n = 30,$$

$$x_1 + x_2 + \dots + x_n = 72.2 + 20.4 + \dots + 55.4 = 999.3$$

Jadi mean nya adalah

$$\bar{x} = \frac{999.3}{30} = 33.1$$

- Median

Rumus untuk mencari median dari sample adalah sebagai berikut

- Jika n ganjil

$$\tilde{x} = x_{(n+1)/2}$$

- Jika n genap

$$\tilde{x} = \frac{1}{2}(x_{n/2} + x_{(n/2)+1})$$

Berbeda dengan mean , pada media disini dibutuhkan data yang terurut dari nilai terkecil hingga terbesar (increasing) agar dapat mengetahui nilai index x seberapa.

1. 8.5	11. 21.5	21. 38.1
2. 8.6	12. 22.3	22. 43.3
3. 9.4	13. 24.1	23. 54.2
4. 12.8	14. 25.1	24. 55.1
5. 13.9	15. 26.2	25. 55.4
6. 14.5	16. 26.5	26. 58.2
7. 17.9	17. 26.5	27. 59.1
8. 19.2	18. 27.3	28. 68.1
9. 20.4	19. 29.1	29. 72.2
10. 20.7	20. 31.9	30. 89.2

Setelah membuat data yang terurut maka kita dapat langsung mencari median nya

Karena nilai n berupa genap maka dapat digunakan formula berikut

$$\tilde{x} = \frac{1}{2}(x_{n/2} + x_{(n/2)+1})$$

$$n = 30, x_{30/2} = x_{15}, x_{(30/2)+1} = x_{16}$$

$$x_{15} = 26.2 \text{ dan } x_{16} = 26.5$$

Maka mediannya adalah

$$\tilde{x} = \frac{1}{2}(x_{15} + x_{16})$$

$$\tilde{x} = \frac{1}{2}(26.2 + 26.5)$$

$$\tilde{x} = \frac{1}{2}(52.7)$$

$$\tilde{x} = 26,35$$

Jadi median dari sample data adalah **26.35**

- Variasi sampel dan standar deviasi

Standard deviasi disimbolkan dengan s adalah nilai dari akar positif dari s^2 dimana s^2 adalah variasi sampel, Jadi untuk itu kita dapat mencari variasi sampel terlebih dahulu agar mendapatkan standar deviasi.

Formula untuk variasi sample adalah:

$$s^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n-1} = \frac{[(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2]}{n-1}$$

Dimana \bar{x} adalah rata rata sampel

Dalam mencari variasi sampel tidak dibutuhkan data yang terurut

Formula untuk standar deviasi adalah:

$$s = \sqrt{s^2}$$

Standar deviasi sampel:

$$n = 30$$

$$\bar{x} = 33.31$$

$$s^2 = \frac{[(8.5-33.31)^2 + (8.6-33.31)^2 + \dots + (72.2-33.31)^2 + (89.2-33.1)^2]}{30-1}$$

$$s^2 = \frac{[447.46024137931]}{29}$$

$$s^2 = 447.46024137931$$

Jadi variasi sampel nya adalah 447.46024137931

Untuk standar deviasinya adalah

$$s = \sqrt{s^2} = \sqrt{447.46024137931}$$

$$s \approx 21.15$$

Jadi standar deviasinya adalah 21.15

- Histogram sampel data

Dalam pembuatan histogram dibutuhkan data yang terurut agar mempermudah untuk mencari range dan interval kelas.

1. 8.5	11. 21.5	21. 38.1
2. 8.6	12. 22.3	22. 43.3
3. 9.4	13. 24.1	23. 54.2
4. 12.8	14. 25.1	24. 55.1
5. 13.9	15. 26.2	25. 55.4
6. 14.5	16. 26.5	26. 58.2
7. 17.9	17. 26.5	27. 59.1
8. 19.2	18. 27.3	28. 68.1
9. 20.4	19. 29.1	29. 72.2
10. 20.7	20. 31.9	30. 89.2

Formula range adalah:

$$Range = x_{\text{terbesar}} - x_{\text{terkecil}}$$

Jadi range nya adalah :

$$x_{\text{terbesar}} = 89.2, x_{\text{terkecil}} = 8.5$$
$$\text{Range} = 89.2 - 8.5 = 80.7$$

Formula lebar kelas (class width)

$$\text{class width} = \frac{\text{Range}}{\text{Jumlah batang histogram}}$$

Disini saya mau memiliki 6 batang pada histogram saya maka class width yang saya peroleh adalah

$$\text{range} = 80.7, \text{jumlah batang histogram} = 6$$

$$\text{class width} = \frac{80.7}{6} = 13.45$$

Biasanya class width dilakukan pembulatan, tapi disini saya tidak akan melakukannya

Lalu buat grup sesuai interval class width, dimulai dari yang terkecil

Classes	Frequency
8.5 - 21.95	
21.95 - 35.4	
35.4 - 48.85	
48.85 - 62.3	
62.3 - 75.75	
75.75 - 89.2	

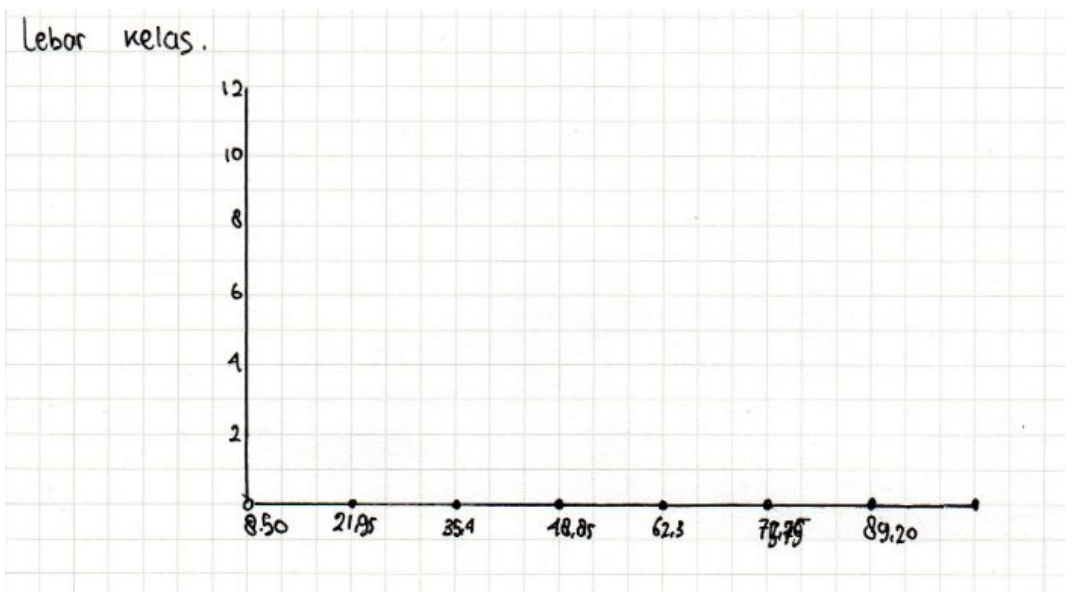
Lalu kita dapat mencari frekuensi data sesuai interval masing masing kelas yang didapatkan

1. 8.5	11. 21.5	21. 38.1
2. 8.6	12. 22.3	22. 43.3
3. 9.4	13. 24.1	23. 54.2
4. 12.8	14. 25.1	24. 55.1
5. 13.9	15. 26.2	25. 55.4
6. 14.5	16. 26.5	26. 58.2
7. 17.9	17. 26.5	27. 59.1
8. 19.2	18. 27.3	28. 68.1
9. 20.4	19. 29.1	29. 72.2
10. 20.7	20. 31.9	30. 89.2

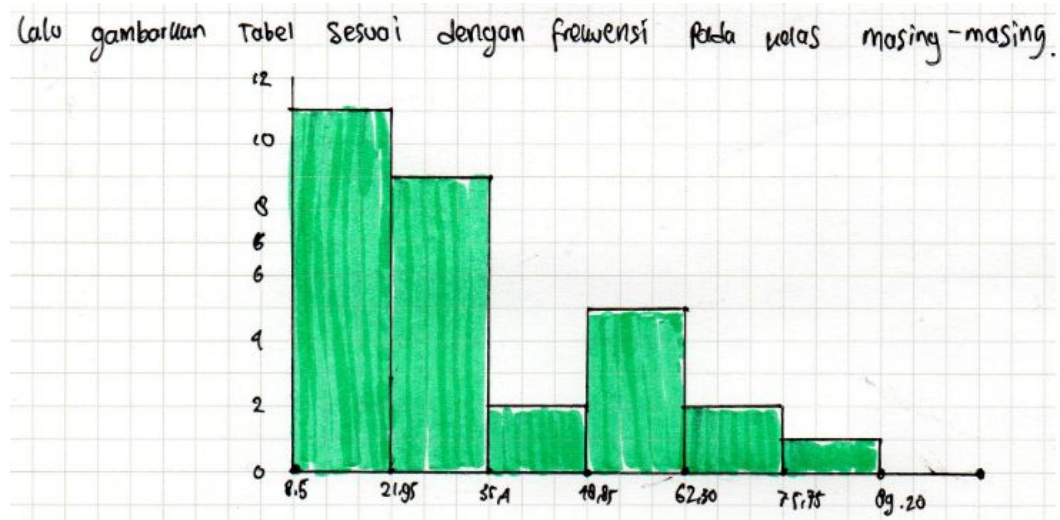
Dapat dilihat frequency pada interval class diatas maka kita dapat melengkapi table yang dibuat

Classes	Frequency
8.5 - 21.95	11
21.95 - 35.4	9
35.4 - 48.85	2
48.85 - 62.3	5
62.3 - 75.75	2
75.75 - 89.2	1

Lalu kita dapat menggambarkan nya dengan garis x adalah kelas dan y adalah frekuensi. Pertama kita menggambarkan lebar kelas terlebih dahulu.



Setelah digambarkan kita dapat menggambarkan tabel sesuai dengan frekuensi kelas masing masing .



2. Diberikan data dibawah berupa masa hidup, dalam jam, dari lima puluh 40Watt, 110-volt lampu pijar yang buram.

1. 919	8. 920	15. 950
2. 1196	9. 948	16. 905
3. 785	10. 1067	17. 972
4. 1126	11. 1092	18. 1035
5. 936	12. 1162	19. 1045
6. 918	13. 1170	20. 855
7. 1156	14. 929	
21. 1195	27. 1237	33. 1009
22. 1195	28. 956	34. 1157
23. 1340	29. 1102	35. 1151
24. 1122	30. 1157	36. 1009
25. 938	31. 978	
26. 970	32. 832	
37. 765	42. 811	47. 1311
38. 958	43. 1217	48. 1037
39. 902	44. 1085	49. 702
40. 1022	45. 896	50. 923
41. 1333	46. 958	

- Rata - rata sampel atau mean

Rumus untuk mencari mean dari sample adalah sebagai berikut

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

dimana n adalah banyaknya data

Dan $x_1 + x_2 + \dots + x_n$ adalah jumlah data

Diketahui :

$$n = 50,$$

$$x_1 + x_2 + \dots + x_n = 919 + 1196 + \dots + 923 = 51354$$

Jadi mean nya adalah

$$\bar{x} = \frac{51354}{50} = 1027.08$$

- Median

Rumus untuk mencari median dari sample adalah sebagai berikut

- Jika n ganjil

$$\tilde{x} = x_{(n+1)/2}$$

- Jika n genap

$$\tilde{x} = \frac{1}{2}(x_{n/2} + x_{(n/2)+1})$$

Berbeda dengan mean , pada media disini dibutuhkan data yang terurut dari nilai terkecil hingga terbesar (increasing) agar dapat mengetahui nilai index x seberapa.

1. 702	15. 936	29. 1037
2. 765	16. 938	30. 1045
3. 785	17. 948	31. 1067
4. 811	18. 950	32. 1085
5. 832	19. 956	33. 1092
6. 855	20. 958	34. 1102
7. 896	21. 958	35. 1122
8. 902	22. 970	36. 1126
9. 905	23. 972	37. 1151
10. 918	24. 978	38. 1156
11. 919	25. 1009	39. 1157
12. 920	26. 1009	40. 1157
13. 923	27. 1022	41. 1162
14. 929	28. 1035	42. 1170

43. 1195	46. 1217	49. 1333
44. 1195	47. 1237	50. 1340
45. 1196	48. 1311	

Setelah membuat data yang terurut maka kita dapat langsung mencari median nya

Karena nilai n berupa genap maka dapat digunakan formula berikut

$$\tilde{x} = \frac{1}{2}(x_{n/2} + x_{(n/2)+1})$$

$$n = 50, x_{50/2} = x_{25}, x_{(50/2)+1} = x_{26}$$

$$x_{25} = 1009 \text{ dan } x_{26} = 1009$$

Maka mediannya adalah

$$\tilde{x} = \frac{1}{2}(x_{25} + x_{26})$$

$$\tilde{x} = \frac{1}{2}(1009 + 1009)$$

$$\tilde{x} = \frac{1}{2}(2018)$$

$$\tilde{x} = 1009$$

Jadi median dari sample data adalah **1009**

- Variasi sampel dan standar deviasi

Standard deviasi disimbolkan dengan s adalah nilai dari akar positif dari s^2 dimana s^2 adalah variasi sampel, Jadi untuk itu kita dapat mencari variasi sampel terlebih dahulu agar mendapatkan standar deviasi.

Formula untuk variasi sample adalah:

$$s^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n-1} = \frac{[(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2]}{n-1}$$

Dimana \bar{x} adalah rata rata sampel

Dalam mencari variasi sampel tidak dibutuhkan data yang terurut

Formula untuk standar deviasi adalah:

$$s = \sqrt{s^2}$$

Standar deviasi sampel:

$$n = 50$$

$$\bar{x} = 1027.08$$

$$s^2 = \frac{[(702-1027.08)^2 + (765-1027.08)^2 + \dots + (1311-1027.08)^2 + (1340-1027.08)^2]}{50-1}$$

$$s^2 = \frac{[1078553.68]}{9}$$

$$s^2 = 22011.299591837$$

Jadi variasi sampel nya adalah 22011.299591837

Untuk standar deviasinya adalah

$$s = \sqrt{s^2} = \sqrt{22011.299591837}$$

$$s \approx 148.36$$

Jadi standar deviasinya adalah 148.36

- **Boxplot**

Untuk menggambar boxplot dibutuhkan beberapa variable antara lain adalah : Q1, Q3, IQR, Median, Oulier range,

Pertama kita mencari median, Q1, Q3, dan IQR

Nilai median telah didapatkan pada pengerjaan soal sebelumnya

$$\bar{x} = 1009$$

Untuk mencari Q1 dan Q3 maka kita mencari median dari upper bound dan lower bound , dapat dilihat dari tabel yang sudah diurutkan bahwa upper bound terdiri dari index x_1 hingga x_{24}

Maka Q1 adalah median dari x_1 hingga x_{24}

$$n = 24$$

Karena nilai n adalah genap maka rumus yang digunakan adalah

$$\tilde{x} = \frac{1}{2}(x_{n/2} + x_{(n/2)+1})$$

$$\tilde{x} = \frac{1}{2}(x_{24/2} + x_{(24/2)+1})$$

$$\tilde{x} = \frac{1}{2}(x_{12} + x_{13})$$

$$x_{12} = 920, x_{13} = 923$$

$$\tilde{x} = \frac{1}{2}(920 + 923)$$

$$\tilde{x} = \frac{1}{2}(1843)$$

$$\tilde{x} = 921.5$$

Jadi Q1 adalah 921.5

Selanjutnya mencari Q3, dimana Q3 adalah median dari x_{26} hingga x_{50}

$$\tilde{x} = \frac{1}{2}(x_{n/2} + x_{(n/2)+1})$$

$$\tilde{x} = \frac{1}{2}(x_{24/2} + x_{(24/2)+1})$$

$$\tilde{x} = \frac{1}{2}(x_{12} + x_{13})$$

*perlu diketahui disini perhitungan index pada Q3 dihitung dari index ke 26

$$x_{12} = 1151, x_{13} = 1156$$

$$\tilde{x} = \frac{1}{2}(1151 + 1156)$$

$$\tilde{x} = \frac{1}{2}(2307)$$

$$\tilde{x} = 1153.5$$

Jadi Q3 bernilai 1153.5

Maka nilai IQR yang didapatkan adalah:

$$IQR = Q_3 - Q_1$$

$$IQR = 1153.5 - 921.5 = 232$$

IQR digunakan untuk mencari range yang **bukan** outlier pada sample data
Dimana range yang bukan outlier adalah

$$[lower\ fence, upper\ fence]$$

Selanjutnya adalah membuat pagar pembatas, upper fence dan lower fence.
Formula untuk mencari upper fence dan lower fence adalah:

$$lower\ fence = Q_1 - 1.5(IQR)$$

$$upper\ fence = Q_3 + 1.5(IQR)$$

Maka lower fence nya adalah:

$$lower\ fence = 921.5 - 1.5(232)$$

$$lower\ fence = 921.5 - 348$$

$$\text{lower fence} = 573.5$$

Jadi lower fence nya adalah 573.5

Dan upper fence nya adalah:

$$\text{upper fence} = 1153.5 + 1.5(232)$$

$$\text{upper fence} = 1153.5 + 348$$

$$\text{upper fence} = 1501$$

Jadi upper fence nya adalah 1501.5

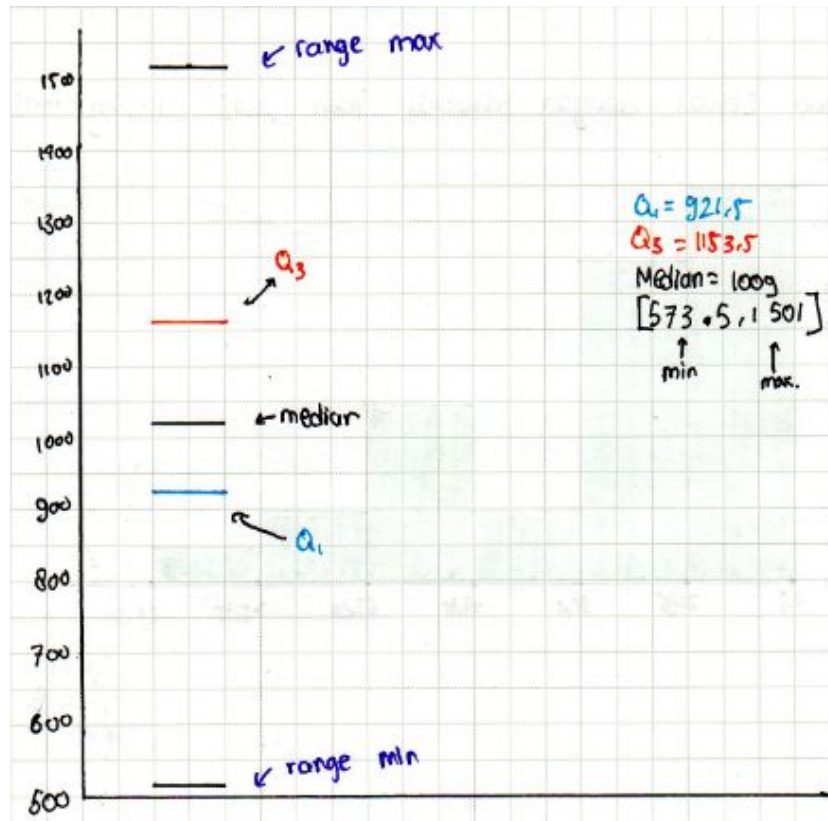
Maka range yang bukan outlier adalah

$$[573.5, 1501]$$

Nilai yang lebih besar atau lebih kecil dari lower dan upper fence dapat dinyatakan sebagai outlier.

Selanjutnya kita dapat menggambarkan boxplot dengan variable - variable yang telah kita hitung sebelumnya

Pertama mari kita gambarkan kerangka nya terlebih dahulu



Lalu setelah itu kita dapat menggambarkan boxplot nya seperti berikut.

